
Dynamics in online social networks

Przemyslaw A. Grabowicz, José J. Ramasco, Víctor M. Eguíluz

Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB),
E07122 Palma de Mallorca, Spain
Email: pms@ifisc.uib-csic.es

1 Introduction

An increasing number of today's social interactions occurs using online social media as communication channels. Some online social networks have become extremely popular in the last decade. They differ among themselves in the character of the service they provide to online users. For instance, Facebook can be seen mainly as a platform for keeping in touch with close friends and relatives, Twitter is used to propagate and receive news, LinkedIn facilitates the maintenance of professional contacts, Flickr gathers amateurs and professionals of photography, etc. Albeit different, all these online platforms share an ingredient that pervades all their applications. There exists an underlying social network that allows their users to keep in touch with each other and helps to engage them in common activities or interactions leading to a better fulfillment of the service's purposes. This is the reason why these platforms share a good number of functionalities, *e.g.*, personal communication channels, broadcasted status updates, easy one-step information sharing, news feeds exposing broadcasted content, etc. As a result, online social networks are an interesting field to study an online social behavior that seems to be generic among the different online services. Since at the bottom of these services lays a network of declared relations and the basic interactions in these platforms tend to be pairwise, a natural methodology for studying these systems is provided by network science. In this chapter we describe some of the results of research studies on the structure, dynamics and social activity in online social networks. We present them in the interdisciplinary context of network science, sociological studies and computer science.

2 Structure of social networks

Social networks in general show a very rich internal structure [1] that in some aspects falls quite far from random graphs or even from artificial networks

created by virtue of a preferential attachment mechanism. In this section we briefly review the most important features broadly found in social networks.

2.1 Degree distribution

The most fundamental characteristic of a network is the distribution of degrees: a function that measures how many friends have the members of the network and what is the variability of this number among all the users. The degree distribution in social networks are usually broad. These distributions have been typically modeled as functions having a heavy tails such as a power-law or log-normals combined with an exponential cut-off at large values of the number of friends [2–7]. This means that there is a large variability in the number of connections of the nodes, with many nodes having small or moderate number of friends and a small number of them maintaining large number of friends. Almost all users of online social networks are connected in a largest connected component [4, 7]. Some of the studies also point out that online social networks contain a densely connected core or cores [4, 5] consisting in groups of high-degree nodes that hold the network together. The existence of such cores provides paths for the connection between distinct parts of the network. A well-known aspect of the social networks is that the average shortest-path distance is low [2, 4, 5, 7]. This characteristic is popularly known as six degrees of separation or small-world effect [8]. The importance of the shortcuts for reducing the network path-length has been highlighted in Ref. [9].

2.2 Triangles and community structure

Possibly, the most important feature distinguishing social networks from other types of networks is their high level of clustering or transitivity [1, 2, 4–7, 9]. The clustering coefficient measures the probability that two nodes sharing a common neighbor (a node to which they are both connected) are connected. This property is quantified with a global clustering coefficient C [6] which is defined as

$$C = \frac{\text{number of closed connected triples}}{\text{number of connected triples}}, \quad (1)$$

where a connected triple of nodes is a sequence of 3 nodes which have at least 2 connections between them, and a closed triple is a triangle. One can also define a local clustering coefficient c_i as:

$$c_i = \frac{\text{number of closed triples centered on node } i}{\text{number of triples centered on node } i}. \quad (2)$$

In this case a global value of clustering coefficient may be obtained averaging the local c_i over all the nodes of the network $\langle c \rangle$. One should note that $\langle c \rangle$ is in general different from the coefficient C , and that the latter has a much

worse scaling behavior. At the structural level, a high clustering coefficient indicates the presence of many triangles in the network. At the social level, this means that friends of an individual tend to be connected between themselves too. This is a well-known phenomena in sociology which is important for the formation of strong social ties [10, 11] and affects to the emergence of positive and negative relations [12]. At the network macroscopic level, a high density of triangles can be related to the existence of community structure in social networks [13]. Furthermore the study [14] suggests that in real networks with high value of clustering coefficient community structure emerges without any additional ingredients included.

Existence of communities in social networks is considered to have high relevance both by sociologist [11, 13] and network scientists [15–17]. We give further arguments on this in the third section of this chapter. In online social networks, groups can be identified in several ways. One of them is searching for communities in the graph defined as more densely connected parts of the network compared with their neighborhood. This approach is usually taken in network science and various community detection algorithms have been developed and continue to be under active development for detecting such clusters [15]. In addition, some online social networks allow their users to create explicit groups and to claim its membership. Although it seems straightforward to make use of such user-declared groups, one should be careful when interpreting them since incentives for creation of such groups may vary [18]. Nevertheless it has been found that declared groups tend to have internally higher clustering coefficient [5] and therefore they may be correlated with the more-densely connected parts of the network found by community detection algorithms.

2.3 Assortativity and homophily

Another common feature of social networks is that connected users tend to be similar [19]. This effect is popularly known as *birds of a feather flock together* phenomena. It manifests itself in social networks through similarities in various properties of connected individuals. From pure network theory point of view the similarity may appear as a correlation of degrees between friends, which is called assortativity mixing, or as a rich-club effect [20]. In such assortative networks nodes of high degrees tend to be connected to other nodes of high degrees, and vice versa, nodes of low degree tend to be connected to other low degree nodes. It has been found that offline social network are assortative in contrast with networks of other types [1, 21]. However, this is not the only property in which friends are similar. This kind of phenomena is in general called homophily and is known to be present very broadly in social networks. People who are connected in online social networks tend to have similar age, live in close locations, and have similar interests [4, 7, 17, 22]. It is also thought that people who belong to the same community, namely the same well-connected group of people, talk about similar topics, which can

have an important impact on information and innovation diffusion in social networks [11, 23].

2.4 Differences between offline and online social networks

As shown in the previous subsections many statistical properties of offline social networks are also found online. On the other hand creating links in a social online network is much less costly than developing offline social relations. These online connections can easily accumulate and pile up to large numbers [24]. If the number of connections increases to the millions, the amount of effort that a user can invest into a relation that each link represents must fall to near zero. An early illustration of the relevance of the definition of social tie in characterization of social networks was shown in the study of email networks: while the distribution of the number of contacts in address books is power law [25], it is exponential when the contacts are restricted to reciprocated emails [26]. Moreover disassortative mixing has been encountered in some online networks [2, 27] in contrast to the assortative mixing characteristic of offline social networks [6]. As a matter of fact there exists an open discussion on the validity of online interactions as indicators of real social activity [24, 28–31]. In order to test the validity of online networks for social studies and to find its limitations further investigation is needed. In this chapter we present some of recent results of such studies.

3 Growth in social networks

3.1 Preferential attachment

Many features of complex systems are characterized by heavy-tailed distributions [32, 33], *e.g.*, frequency of words [34], the wealth of nations [35], degree distribution of complex networks [36], etc. This property is typically perceived as a symptom of the rich-gets-richer principle, from which the so-called preferential growth models stem. The common concept of these models is that the elements of the system grow proportionally to their current size, what is referred to as preferential growth or preferential attachment rule. Typically, in these models, increments of the defining variables of the system occur in each time step. Such increments can involve the addition of new elements and/or to increase the sizes of the existing ones according to a preferential growth rule. Preferential models are usually the first approach to explain heavy-tailed distributions in many different systems [37–40]. In the case of networks, this kind of models got popularized a decade ago [36, 41–44]. The first of these models in the context of complex networks was introduced by Barabási and Albert in Ref. [41]. To describe it shortly: in each time step one node is introduced to the system with m edges. These edges are connected to existing nodes in the system with probability proportional to the degree of the present nodes. As

a result a network with heavy-tailed (usually power-law) distribution of node degrees emerges. The rules of Barabasi's model yield high simplicity, which is typically a desirable feature, but that can be too rigid in some cases. In preferential-growth models the time unit is directly coupled to the number of new arriving elements, which can complicate the comparison of the dynamics of these models with real data. Some other drawbacks include the lack of heterogeneity and strong correlation between age of elements and their size [45]. Because of these issues the basic preferential growth model is typically used as a simple toy-model for generation of networks with power-law distribution of degrees. On the other hand, it is also successfully used as a component of models trying to simulate growth of real social networks [46, 47].

3.2 Heterogeneity

In many real systems, especially in social systems, individuals or elements are very diverse. This factor is related to the heavy-tailed distributions that are so commonly found. In this direction, some models incorporating heterogeneity in the form of fitness, hidden variables or ranking shuffling have been proposed [48–52]. In general this family of models determines growth of elements with some kind of intrinsic property. Whereas in preferential attachment models the growth was proportional to current size of the elements, in fitness models it is usually proportional to the intrinsic fitness of each element. Typically the fitness is a random variable specific for each element drawn from a given variables. This introduces high heterogeneity between the elements. A number of empirical works shows how this intrinsic fitness is distributed and what is its role in complex system growth [53–56]. We discuss in detail one of the models of this family in the next section when commenting on the growth of groups in online social networks.

3.3 Triadic closure / Triangle closing

Due to the fact that clustering coefficient is remarkably high in some networks (mainly social networks), other growth models have been introduced in order to reproduce high number of triangles in the network. One of the first models accounting for this was [9] in which regular network with initially high clustering had its connections rewired to make it more realistic and control clustering coefficient, as well as average shortest path length. A more sophisticated model used to simulate growth of social networks has been proposed in Ref. [47] and one of its main components is triangle closing. In this model new nodes appearing in the system connect to some node, usually using preferential attachment rule, and then start closing triangles with neighbors of this node. This simple triangle closing mechanism exhibit much more realistic results in modeling online social networks [47].

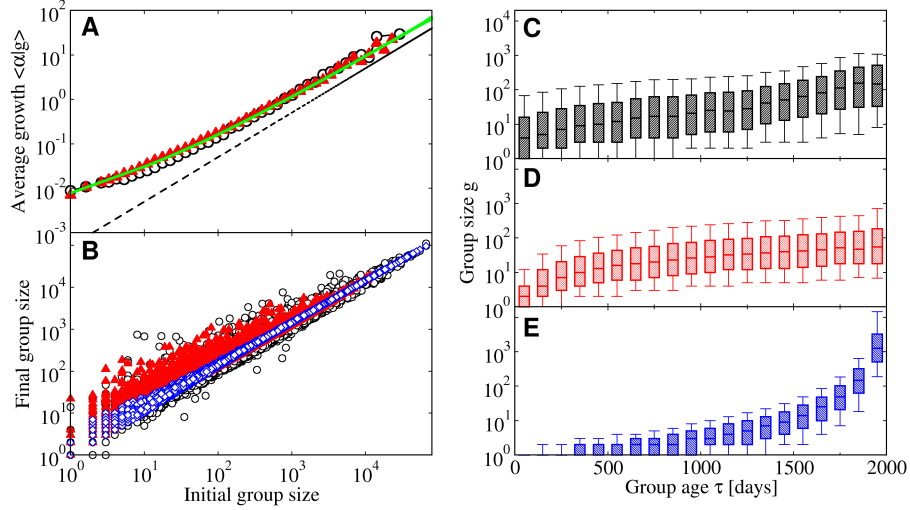


Fig. 1. Growth of groups in Flickr. (A) Average daily growth as a function of the initial size of the groups, estimated for the period of 6 weeks and averaged over 260,000 groups of a given initial size, for: the real data from Flickr (circles), the heterogeneous linear growth model (triangles). The dashed line corresponds to the linear behavior $\langle \alpha | g \rangle \sim g$. (B) Initial and final group sizes over a period of 350 days for the real data (circles), the heterogeneous linear growth model (filled triangles) and Simon model (diamonds). Each point represents a single group, there are 9,503 points plotted for each set of points. (C-E) Box plots with whiskers at 9th/91st percentile of final size of groups as a function of their age at the time of the measurement for 260,000 groups for (C) the real data, (D) the heterogeneous linear growth model, (E) the Simon model. Adapted from Ref. [56].

3.4 Dynamics of groups

As we have emphasized in the previous section the existence of communities plays an important role in functioning of social networks. In this section we present studies of the growth of such groups. Several aspects have been identified as positively influencing groups' growth and their persistence. It has been suggested that high internal connectivity helps declared groups' growth [57]. Other work argues that flexibility of big communities helps them stay alive longer, while small groups are more persistent if their composition stays unchanged [17].

From the macroscopic perspective growth of groups can be described and modeled using a version of preferential attachment model or a model with hidden variables/intrinsic heterogeneity. A comparison between these two approaches has been performed in Ref. [56] using real data from Flickr. The heterogeneous linear growth model suggested in this study assumes linear growth of groups with growth value (fitness) being drawn from heavy-tailed

distribution (log-normal) and a number of new groups appearing in the systems growing linearly in time. As a comparison, a version of Simon model [37] has been used, which represents a model from preferential attachment family. As one can see in Figure 1A, the average growth $\langle \alpha|g \rangle$ for groups of given size g is proportional to the size of the groups for high g . This commonly is interpreted as the consequence of preferential attachment. However, as it is shown in Figure 1A, one obtains similar dependence using the heterogeneous linear growth model. This is the case because the average growth is an average over all groups of a given size, each of them growing linearly. Due to the heterogeneity and the linear growth, at a given time larger groups consist of old groups that grow slowly and younger groups that grow faster. Thus, the observation of preferential growth for groups of the same size does not reflect in this case an underlying rich-gets-richer principle, but it is a consequence of the competition of groups with different growth values and ages. Both the heterogeneous linear growth model and the Simon model produce heavy-tailed distribution of group sizes. However, the former model performs better in other respects. First, in the Simon model the final size of groups is heavily determined by their initial size measured one year before (Figure 1B), thus there is little heterogeneity among the groups, in contrast to the heterogeneous linear growth model which displays a degree of heterogeneity similar to the one of real groups. Second, for the Simon model the correlation of size and age is strong, while it is weak for real groups and the heterogeneous linear growth model (Figure 1C-E). In the heterogeneous linear growth model the heavy-tailed distribution of final sizes of elements does not emerge from the growth process itself (e.g., rich-gets-richer principle), but from the intrinsic heterogeneity of elements which take part in this growth process. This certainly does not answer the question why some groups grow faster than the others, as we do not understand yet what factors influence the fitness of the groups. However it points that it does not have to be due to the fact that one group is bigger than the other as in preferential attachment models. The simplicity of this approach suggests that the characterization of the heterogeneity may play an important role in understanding the origin of broad distributions and the time evolution of many real systems.

4 Activity in the online social networks

In general a social network is a broad term and it refers to a set of actors and a set of ties between them representing some kind of relation or interaction. In fact, however, there are many types of both relations and interactions, and usually they happen on top of each other. So far we mostly discussed social networks which represent a particular relation or interaction, e.g.: co-appearance in movies, boards of directors or co-authorship [1, 6, 21], network of online friendship [2, 5, 7, 47], network of communication [4, 17, 58]. In online social networks, we can relate user activity with their declared relations with

other users. In other words, one can relate pairwise (rarely one-to-many) interactions of users with their declared social network. We describe the studies which tackle this issue in along the section.

4.1 Activity networks versus declared social network

The comparison of the network built from declared online friends and the network built from user interactions shows several differences at the structural level. First of all, the actors tend to interact with much less people than they declare as friends, what results in smaller degrees of nodes in the interactions network [59, 60]. Moreover, the friends they interact with change rapidly and only about 30% of pairwise interactions in one month continue over the next month [59]. Due to the fact that the degrees are lower, the properties related to small-world effect are also less evident, namely average path lengths are higher [60] and there is less densely connected cores [61].

4.2 Theories on social ties and information diffusion

The theory known as *the strength of weak ties*, proposed by Granovetter [11], deals with the relation between structure, intensity of social ties and diffusion of information in offline social networks. On one hand, a tie can be characterized by its strength, which is related to the time spend together, intimacy and emotional intensity of a relation. Strong ties refer to relations with close friends or relatives, while weak ties represent links with distant acquaintances. On the other hand, a tie can be characterized by its position in the network. Social networks are usually composed of communities. A tie can thus be internal to a group or a bridge between groups, as in Figure 2. Granovetter's theory predicts that weak ties act as bridges between groups and are important for the diffusion of new information across the network. Strong ties are predicted to be located at the interior of the groups between actors who have many friends in common. Burt's work [13] emphasizes the advantage of connecting different groups to access novel information due to the diversity in the sources.

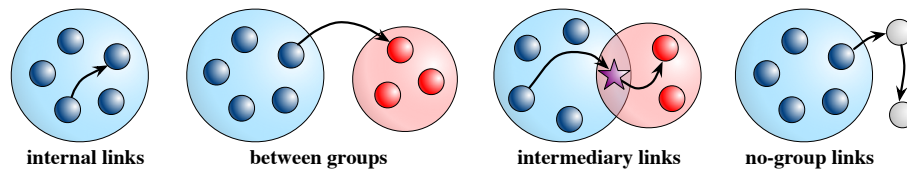


Fig. 2. Different types of links depending on their position with respect to the groups' structure: internal, between groups, intermediary links and no-group links. Adapted from Ref. [16].

Furthermore more recent works point out that information propagation may be dependent on the type of content transmitted [62–64] and on a *diversity-bandwidth tradeoff* [65]. The bandwidth of a tie is defined as the rate of information transmission per unit of time. Aral et al. [65] note that weak ties interact infrequently, therefore have low bandwidth, whereas strong ties interact more often and have high bandwidth. The authors claim that both diversity and bandwidth are relevant for the diffusion of novel information. Since both are anticorrelated, there has to be a tradeoff to reach an optimal point in the propagation of new information. They also suggest that strong ties may be important to propagate information depending on the structural diversity, the number of topics and the dynamic of the information. Due to the different nature of online and offline interactions, it is not clear whether online networks organize following the sociological theories. In the following subsection we present results of some works testing if these theories apply to online social networks.

4.3 Testing social theories in online social networks

The predictions of the theory of *the strength of weak ties* have been checked in a mobile phone calls dataset [58] and, very recently, in online social networks [16, 66, 67]. Different predictors have been considered to estimate social tie strength [68] including, for instance, time spent together [68], the duration of phone calls [58] or number of messages exchanged [16, 66]. The two works [16, 58] have measured the dependence of strength of a tie on number of common friends shared by the two actors, showing that the more friends they share the more likely it is that the tie is strong. This stays in agreement with homophily effect in social network described at the beginning of this chapter. Many shared friends of a pair of users coupled by a strong tie can be interpreted as high homophily between them in terms of acquired friends. Furthermore, large field experiment performed at Facebook [66] has isolated the effect of homophily and social impact on the probability of propagation of information in online social network. The study has shown that users are around 7-times more likely to re-broadcast a piece of information published by their friends if they are exposed to it, which is interpreted as 7-times higher chance of information propagation due to social influence than to homophily. Moreover, the work argues that the weaker is the tie for which information propagation is considered, the higher is the likelihood of information flow due to social influence. This corresponds to Granovetter’s prediction that weak ties are important for information diffusion. In the following paragraphs we describe in more detail findings of a similar study in Twitter [16], a popular social microblogging platform.

Online networks are promising for social studies due to the wide availability of data and the fact that different types of interactions are explicitly separated: e.g., information diffusion events are distinguished from more personal communications. Diffusion events are implemented as a system option

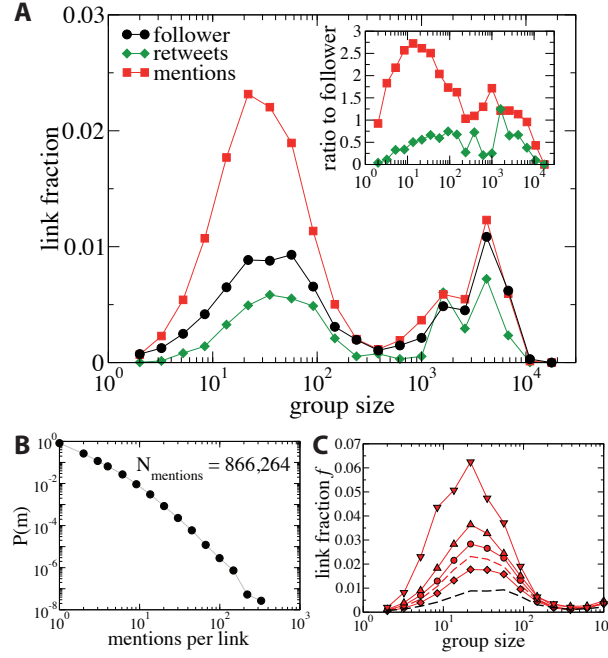


Fig. 3. Internal activity in Twitter. (A) Fraction f of internal links as a function of the group size in number of users. The curve for the follower network acts as baseline (black) for mentions (red) and retweets (green). Note that if mentions/retweets were randomly appearing over follower links then the red/green curve should collapse with the black curve. Inset: links with mentions divided by the baseline (red), and links with retweets divided by the baseline (green). (B) Distribution of the number of mentions per link. (C) Fraction of links with mentions as a function of their intensity. The dashed curves are the total for the follower network (black) and for the links with mentions (red). While the other curves correspond (from bottom to top) to fractions of links with: 1 non-reciprocated mention (diamonds), 3 mentions (circles), 6 mentions (triangle up) and more than 6 reciprocated mentions (triangle down). From Ref. [16].

in the form of *share*, or *repost* buttons with which it is enough to single-click on a piece of information to rebroadcast it to all the users' contacts. This is in contrast to personal communications for which more effort has to be invested to write a short message and to select the recipient(s). In Twitter such actions are called respectively *retweet* [69] and *mention/reply* [70]. The more mentions has been exchanged between two users, even more so if reciprocated, the stronger we consider the tie between them. On the other hand declared network does also exist in Twitter and is made of directed follower links. One, using clustering algorithms, can find communities of more densely connected users in such network. Specifically, in the study which we present,

various clustering algorithms have been used [71] and for brevity we will focus only on results for OSLOM [72]. Granovetter theory predicts that social ties should occur more often inside communities. This is what happens for links with mentions. We define the fraction f as ratio between the number of links with specified type of interaction in given position with respect to the groups of corresponding size and the total number of links with that interaction. The fractions f reveals an interesting pattern as function of the group size as can be seen in Figure 3A. Links with mentions are more abundant inside communities than any other links. This effect is especially significant for groups of sizes from 10 to 150 members. In addition, the distribution of the number of times that a link is used (intensity) for mentions is wide, which allows for a systematic study of the dependence of intensity and position (see Figure 3B). It turns out that the more intense (or reciprocated) a link with mentions is, the more likely it becomes to find this link as internal (Figure 3C). This corresponds to Granovetter expectation that the stronger the tie is the higher the number of mutual contacts of both parties it has and the higher the chance that the parties belong to the same group.

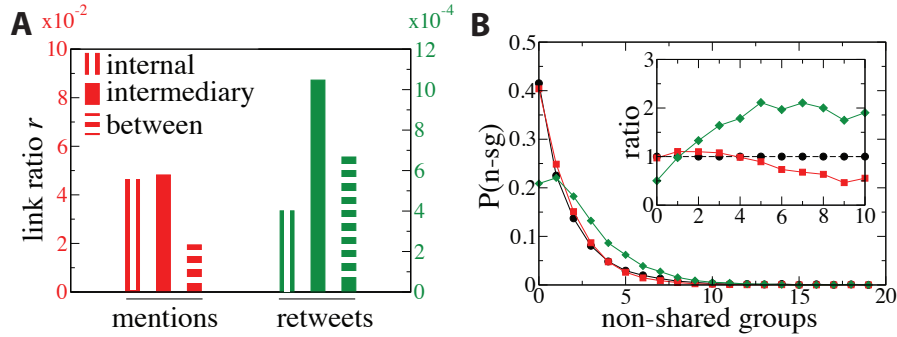


Fig. 4. Intermediary links. (A) Ratio r between the number of links with mentions or retweets and number of follower links. (B) Distribution of the links in the follower network (black curve), those with mentions (red curve) and retweets (green curve) as a function of the number of non-shared groups of the users connected by the link. Inset, ratios between these distributions and the follower network. From Ref. [16].

The communication between groups can take place in two ways: the information can propagate by means of links between groups or by passing through an intermediary user belonging to more than one group, see Figure 2. We have defined as intermediary the links connecting a pair of users sharing a common group and with at least one of the users belonging also to a different group. In order to estimate the efficiency of the different types of links as attractors of mentions and retweets, there was measured a ratio r defined as number of links with specified interaction in a given position divided by the total num-

ber of links in that position. The bar plot with the values of r is displayed in Figure 4. The efficiency of the different type of links can thus be compared for the attraction of mentions (red bars) and retweets (green bars). Links internal to the groups attract more mentions and less retweets than links between groups in agreement with the predictions of the strength of weak ties theory. Intermediary links attract mentions as likely as internal links: the ratio of intermediary links with mentions is very close to the ratio of internal links with mentions. This is expected because intermediary links are also internal to the groups. However, the aspect that differentiates more intermediary links from other type of links is the way that they attract retweets. Intermediary links bear retweets with a higher likelihood than either internal or between-groups connections (see Figure 4A). This fact can be interpreted within the framework of the tradeoff between diversity and bandwidth [65]: strong ties are expected to be internal to the groups and to have high bandwidth, while ties connecting diverse environments or groups are more likely to propagate new information. High bandwidth links in our case correspond to those with multiple mentions, while links providing large diversity are the ones between groups. Intermediary links exhibit these two features: they are internal to the groups and statistically bear more mentions, and introduce diversity through the intermediary user membership in several groups. Moreover, in line with the theories [11, 13, 65], higher diversity increases the chances for a link to bear retweets as can be seen in Figure 4B, which implies a more efficient information flow. In the inset of the Figure it is shown that the number of non-shared groups assigned to the users connected by the link positively correlates with, up to twice higher than expected, number of retweets.

5 Summary

Research of online social networks is a rich and an active field of study. The availability of large amount of data allows for studies of both dynamics of social networks and user-user activity on the social network connections. Different growth models have been proposed to simulate the growth of the network, among which three main families are: preferential growth models, fitness or hidden variables models and triangle-closing models. The latter model is reported to yield most accurate results; however, it also incorporates mechanism of preferential attachment. The main advantage of triangle-closing model is that it directly produces network with enough clustering, which is reported to be a feature of social networks. Moreover, there are still open questions about the origin of these mechanisms and of some other phenomena observed during the growth process such as network densification [73]. While declared social network evolves different types of interactions occur between its members, mostly between users already connected in the declared social network.

Recent studies have shown that different types of interactions happen according to the patterns predicted by the sociological theories. In general strong

ties, which in online social networks are usually defined as the links with many messages exchanged between the pair of users, happen more often between users who have many friends in common, or who belong to the same communities. On the other hand, weak ties appear more often between users who do not share friends and belong to different groups. It has been shown that weak ties are more efficient for the information spreading than strong ties. Closer study shows that tradeoff between diversity and bandwidth may be crucial for diffusion of information.

In conclusion dynamics and activity in online social networks is remarkably rich and tells us much about our social behavior and confirms some of the known offline social theories. We expect that this field of research will be active and developing in the following years and that numerous further online observations and experiments will be undertaken to better understand and quantitatively describe social behaviors.

6 Acknowledgments

We acknowledge partial support from the European Commission through PATRES project, the Spanish Ministry of Economy (MINECO) and FEDER (EU) through projects FISICOS (FIS2007-60327) and MODASS (FIS2011-247852). P.A.G. acknowledges support from the JAE program of the CSIC; J.J.R. acknowledges support from the Ramón y Cajal program of MINECO.

References

- [1] M. NEWMAN and J. PARK, *Phys. Rev. E* **68**, 9 (2003).
- [2] Y. Y. AHN, S. HAN, H. KWAK, S. MOON, and H. JEONG, Analysis of topological characteristics of huge online social networking services, in *Proceedings of the 16th international conference on World Wide Web - WWW '07*, pp. 835–844, Banff, Alberta, Canada, 2007, ACM.
- [3] S. BOCCALETTI, V. LATORA, Y. MORENO, M. CHAVEZ, and D. HWANG, *Physics Reports* **424**, 175 (2006).
- [4] J. LESKOVEC and E. HORVITZ, Planetary-scale views on a large instant-messaging network, in *Proceeding of the 17th international conference on World Wide Web*, pp. 915–924, ACM, 2008.
- [5] A. MISLOVE, M. MARCON, K. P. GUMMADI, P. DRUSCHEL, and B. BHATTACHARJEE, Measurement and analysis of online social networks, in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement - IMC '07*, pp. 29–42, San Diego, California, USA, 2007, ACM.
- [6] M. E. J. NEWMAN, *SIAM review* **45**, 167 (2003).
- [7] J. UGANDER, B. KARRER, L. BACKSTROM, and C. MARLOW, *Arxiv preprint arXiv:1111.4503* (2011).

- [8] J. TRAVERS and S. MILGRAM, *Sociometry* **4**, 425 (1969).
- [9] D. J. WATTS and S. H. STROGATZ, *Nature* **393**, 440 (1998).
- [10] D. KRACKHARDT and M. HANDCOCK, Heider vs Simmel: Emergent Features in Dynamic Structures, in *Statistical Network Analysis: Models, Issues, and New Directions*, edited by E. AIROLDI, D. BLEI, S. FIENBERG, A. GOLDENBERG, E. XING, and A. ZHENG, volume 4503 of *Lecture Notes in Computer Science*, pp. 14–27, Springer Berlin / Heidelberg, 2007.
- [11] M. S. GRANOVETTER, *American Journal of Sociology* **78**, 1360 (1973).
- [12] J. LESKOVEC, D. HUTTENLOCHER, and J. KLEINBERG, Predicting positive and negative links in online social networks, in *Proceedings of the 19th international conference on World wide web*, pp. 641–650, New York, New York, USA, 2010, ACM.
- [13] R. BURT, *Brokerage and closure: An introduction to social capital*, Oxford University Press, USA, 2005.
- [14] D. FOSTER, J. FOSTER, P. GRASSBERGER, and M. PACZUSKI, *Phys. Rev. E* **84**, 066117 (2011).
- [15] S. FORTUNATO, *Physics Reports* **486**, 75 (2010).
- [16] P. A. GRABOWICZ, J. J. RAMASCO, E. MORO, J. M. PUJOL, and V. M. EGUILUZ, *PLoS ONE* **7**, e29358 (2012).
- [17] G. PALLA, A.-L. BARABÁSI, and T. VICSEK, *Nature* **446**, 664 (2007).
- [18] N. PISSARD and C. PRIEUR, Thematic vs. social networks in web 2.0 communities: A case study on Flickr groups, in *Proc. of Algotel Conference*, 2007.
- [19] M. MCPHERSON, L. SMITH-LOVIN, and J. M. COOK, *Annual Review of Sociology* **27**, 415 (2001).
- [20] T. OPSAHL, V. COLIZZA, P. PANZARASA, and J. RAMASCO, *Phys. Rev. Lett.* **101**, 168702 (2008).
- [21] M. NEWMAN, *Phys. Rev. Lett.* **89**, 208701 (2002).
- [22] R. SCHIFANELLA, A. BARRAT, C. CATTUTO, B. MARKINES, and F. MENCZER, Folks in Folksonomies, in *Proceedings of the third ACM international conference on Web search and data mining - WSDM '10*, p. 271, New York, New York, USA, 2010, ACM Press.
- [23] D. CENTOLA, *Science* **334**, 1269 (2011).
- [24] A. AVNIT, <http://tinyurl.com/nshcjg> (2009).
- [25] H. EBEL, L.-I. MIELSCH, and S. BORNHOLDT, *Phys. Rev. E* **66**, 035103 (2002).
- [26] R. GUIMERÀ, L. DANON, A. DÍAZ-GUILERA, F. GIRALT, and A. ARENAS, *Phys. Rev. E* **68**, 065103 (2003).
- [27] H.-B. HU and X.-F. WANG, *Europhys. Lett.* **86**, 18003 (2009).
- [28] J. N. CUMMINGS, B. BUTLER, and R. KRAUT, *Communications of the ACM* **45**, 103 (2002).
- [29] D. LAZER, A. PENTLAND, L. ADAMIC, S. ARAL, A.-L. BARABASI, D. BREWER, N. CHRISTAKIS, N. CONTRACTOR, J. FOWLER, M. GUTMANN, T. JEBARA, G. KING, M. MACY, D. ROY, and M. VAN ALSTYNE, *Science* **323**, 721 (2009).

- [30] A. VESPIGNANI, *Science* **325**, 425 (2009).
- [31] D. J. WATTS, *Nature* **445**, 489 (2007).
- [32] M. NEWMAN, *Contemporary Physics* **46**, 323 (2005).
- [33] A. SAICHEV, Y. MALEVERGNE, and D. SORNETTE, *Theory of Zipf's Law and Beyond*, Springer, New York, 2009.
- [34] G. ZIPF, *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*, Addison-Wesley Press, Reading MA, 1949.
- [35] V. PARETO, *Cours d'Économie Politique*, Librairie Droz, Geneve, 1964.
- [36] A.-L. BARABÁSI and R. ALBERT, *Science* **286**, 509 (1999).
- [37] H. A. SIMON, *Biometrika* **42**, 425 (1955).
- [38] E. EISENBERG and E. LEVANON, *Phys. Rev. Lett.* **91**, 138701 (2003).
- [39] K. YAMASAKI, K. MATIA, S. BULDYREV, D. FU, F. PAMMOLLI, M. RICCABONI, and H. STANLEY, *Phys. Rev. E* **74**, 035103 (2006).
- [40] Y. E. MARUVKA, D. A. KESSLER, and N. M. SHNERB, *PLoS ONE* **6**, e26480 (2011).
- [41] A.-L. BARABÁSI, R. ALBERT, and H. JEONG, *Physica A* **272**, 173 (1999).
- [42] B. A. HUBERMAN and L. A. ADAMIC, *Nature* **401**, 131 (1999).
- [43] S. N. DOROGVTSEV, J. F. F. MENDES, and A. N. SAMUKHIN, *Phys. Rev. Lett.* **85**, 4633 (2000).
- [44] S. BORNHOLDT and H. EBEL, *Phys. Rev. E* **64**, 035104 (2001).
- [45] L. A. ADAMIC and B. A. HUBERMAN, *Science* **287**, 2115 (2000).
- [46] A. MISLOVE, H. S. KOPPULA, K. P. GUMMADI, P. DRUSCHEL, and B. BHATTACHARJEE, Growth of the flickr social network, in *Proceedings of the first workshop on Online social networks - WOSP '08*, pp. 25–30, Seattle, WA, USA, 2008, ACM.
- [47] J. LESKOVEC, L. BACKSTROM, R. KUMAR, and A. TOMKINS, Microscopic evolution of social networks, in *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '08*, pp. 462–470, Las Vegas, Nevada, USA, 2008, ACM.
- [48] G. BIANCONI and A.-L. BARABÁSI, *Europhys. Lett.* **54**, 436 (2001).
- [49] G. CALDARELLI, A. CAPOCCI, P. DE LOS RIOS, and M. A. MUÑOZ, *Phys. Rev. Lett.* **89**, 258702 (2002).
- [50] B. SÖDERBERG, *Phys. Rev. E* **66**, 066121 (2002).
- [51] M. BOGUÑÁ and R. PASTOR-SATORRAS, *Phys. Rev. E* **68**, 036112 (2003).
- [52] J. RATKIEWICZ, S. FORTUNATO, A. FLAMMINI, F. MENCZER, and A. VESPIGNANI, *Phys. Rev. Lett.* **105**, 158701 (2010).
- [53] D. GARLASCHELLI and M. LOFFREDO, *Phys. Rev. Lett.* **93**, 188701 (2004).
- [54] G. DE MASI, G. IORI, and G. CALDARELLI, *Phys. Rev. E* **74**, 066112 (2006).
- [55] J. S. KONG, N. SARSHAR, and V. P. ROYCHOWDHURY, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13724 (2008).

- [56] P. A. GRABOWICZ and V. M. EGUÍLUZ, *Europhys. Lett.* **97**, 28002 (2012).
- [57] D. TARABORELLI, Viable web communities: two case studies, in *Viability and Resilience of Complex Systems*, edited by G. DEFFUANT and N. GILBERT, pp. 75–105, Kluwer, 2011.
- [58] J.-P. ONNELA, J. SARAMÄKI, J. HYVÖNEN, G. SZABÓ, D. LAZER, K. KASKI, J. KERTÉSZ, and A.-L. BARABÁSI, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7332 (2007).
- [59] B. VISWANATH, A. MISLOVE, M. CHA, and K. P. GUMMADI, On the Evolution of User Interaction in Facebook, in *Proceedings of the 2nd ACM workshop on Online social networks - WOSN '09*, p. 37, Barcelona, Spain, 2009, ACM Press.
- [60] C. WILSON, B. BOE, A. SALA, K. P. PUTTASWAMY, and B. Y. ZHAO, User interactions in social networks and their implications, in *Proceedings of the fourth ACM european conference on Computer systems - EuroSys '09*, pp. 205–218, Nuremberg, Germany, 2009, ACM Press.
- [61] H. CHUN, H. KWAK, Y. H. EOM, Y. Y. AHN, S. MOON, and H. JEONG, Comparison of online social relations in volume vs interaction: a case study of cyworld, in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement conference - IMC '08*, pp. 57–70, Vouliagmeni, Greece, 2008, ACM.
- [62] D. CENTOLA, V. M. EGUÍLUZ, and M. W. MACY, *Physica A* **374**, 449 (2007).
- [63] D. CENTOLA and M. MACY, *American Journal of Sociology* **113**, 702 (2007).
- [64] D. CENTOLA, *Science* **329**, 1194 (2010).
- [65] S. ARAL and M. VAN ALSTYNE, *American Journal of Sociology* **117**, 90 (2011).
- [66] E. BAKSHY, I. ROSENN, C. MARLOW, and L. ADAMIC, The role of social networks in information diffusion, in *Proceedings of the 21st international conference on World Wide Web - WWW '12*, p. 519, New York, New York, USA, 2012, ACM Press.
- [67] J. L. IRIBARREN and E. MORO, *Social Networks* **33**, 134 (2011).
- [68] P. V. MARSDEN and K. E. CAMPBELL, *Social Forces* **63**, 482 (1984).
- [69] W. GALUBA, K. ABERER, D. CHAKRABORTY, Z. DESPOTOVIC, and W. KELLERER, Outtweeting the twitterers-predicting information cascades in microblogs, in *Proceedings of the 3rd conference on Online social networks*, USENIX Association, 2010.
- [70] C. HONEYCUTT and S. HERRING, Beyond microblogging: Conversation and collaboration via Twitter, in *42st Hawaii International International Conference on Systems Science*, edited by N. FIELDING, R. M. LEE, and G. BLANK, pp. 1–10, Waikoloa, Big Island, HI, USA, 2009, IEEE.
- [71] P. A. GRABOWICZ, J. J. RAMASCO, E. MORO, J. M. PUJOL, and V. M. EGUÍLUZ, <http://www.scribd.com/doc/99259702/Supporting->

- Information-to-Social-Features-of-Online-Networks-The-Strength-of-Intermediary-Ties-in-Online-Social-Media* (2012).
- [72] A. LANCICHINETTI, F. RADICCHI, J. J. RAMASCO, and S. FORTUNATO, *PLoS ONE* **6**, e18961 (2011).
- [73] J. LESKOVEC, J. KLEINBERG, and C. FALOUTSOS, *ACM Transactions on Knowledge Discovery from Data* **1**, Article 2 (2007).